REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
08-04-2010	Technical Paper	

4. TITLE AND SUBTITLE

QSPR and Artificial Neural Network Predictions of Hypergolic Ignition Delays for Energetic Ionic Liquids

5a. CONTRACT NUMBER
FA8650-09-M-2037

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Debasis Sengupta, Maciej Z. Pindera, & J. Vernon Cole (CFD Research Corporation)

5d. PROJECT NUMBER

5f. WORK UNIT NUMBER
300509WG

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

CFD Research Corporation
215 Wynn Drive
Huntsville AL 35805

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-RZ-ED-TP-2010-137

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory (AFMC)
AFRL/RZS
5 Pollux Drive
Edwards AFB CA 93524-7048

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S NUMBER(S)
AFRL-RZ-ED-TP-2010-137

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited (PA #10166).

13. SUPPLEMENTARY NOTES
For presentation at the 57th JANNAF Joint Subcommittee Meeting, Colorado Springs, CO, 3-7 May 2010.

14. ABSTRACT

Due to their negligible volatility, energetic ionic liquids are being considered as next generation hypergolic fuels for replacing toxic monomethylhydrazine. One design challenge for energetic ionic liquids is to maintain their ignition delays as close to that of monomethylhydrazine. The ignition process of ionic liquids with an oxidizer, such as nitric acid, is a complex process and, to date, there is no theoretical method for predicting the ignition delay. The present work examined two correlation methods, Quantitative Structure Property Relationship (QSPR) and Artificial Neural Networks (ANNs), for their ability to predict this quantity. A set of five descriptors were chosen from a pool of more than 160 to establish these correlations. A good QSPR correlation was obtained using these descriptors. We also trained an artificial neural network and examined the predictive ability of the network using an extensive 5-fold cross validation process for the same set of descriptors. A number of data normalization techniques were examined for network training and validation. The results show that ANNs exhibit excellent prediction capabilities for this application.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Ghanshyam Vaghjiani
Unclassified	**Unclassified**	**Unclassified**	SAR	17	19b. TELEPHONE NUMBER *(include area code)* N/A

QSPR AND ARTIFICIAL NEURAL NETWORK PREDICTIONS OF HYPERGOLIC IGNITION DELAYS FOR ENERGETIC IONIC LIQUIDS

Debasis Sengupta*, Maciej Z. Pindera, J. Vernon Cole
CFD Research Corporation
215 Wynn Dr., Huntsville, AL-35805

Ghanshyam L. Vaghjiani
Propellants Branch
Space and Missile Propulsion Division,
Propulsion Directorate, Air Force Research Laboratory, AFRL/RZSP
10 E Saturn Boulevard, Edwards AFB, California 93524

ABSTRACT

Due to their negligible volatility, energetic ionic liquids are being considered as next generation hypergolic fuels for replacing toxic monomethylhydrazine. One design challenge for energetic ionic liquids is to maintain their ignition delays as close to that of monomethylhydrazine. The ignition process of ionic liquids with an oxidizer, such as nitric acid, is a complex process and, to date, there is no theoretical method for predicting the ignition delay. The present work examined two correlation methods, Quantitative Structure Property Relationship (QSPR) and Artificial Neural Networks (ANNs), for their ability to predict this quantity. A set of five descriptors were chosen from a pool of more than 160 to establish these correlations. A good QSPR correlation was obtained using these descriptors. We also trained an artificial neural network and examined the predictive ability of the network using an extensive 5-fold cross validation process for the same set of descriptors. A number of data normalization techniques were examined for network training and validation. The results show that ANNs exhibit excellent prediction capabilities for this application.

INTRODUCTION

Hypergolic bipropellants have been employed for many tactical and small-to-medium rocket engines since World War II. Hypergolic fuels react nearly instantaneously when in contact with oxidizers, such as various forms of nitric acid. A major advantage of a hypergolic propulsion system is that it does not require a separate ignition system. Although monomethylhydrazine (MMH) has been used for decades, it is highly toxic and carcinogenic. This fuel also has high vapor pressure (volatility) at room temperature, which makes it difficult to handle and transport. In addition, recent environmental and health concerns have encouraged the development of environmentally-friendly fuels. Therefore, one of the thrust areas in fuel research is the development of hypergolic propellants that can replace MMH without sacrificing performance. Recently Energetic Ionic Liquids (EILs) have emerged as an attractive alternative due to their potential applications as hypergolic bipropellants[1-5]. Due to their negligible vapor pressure (volatility), they are more environment-friendly and safe to handle and transport than the currently used hydrazine based fuels. In addition, EILs are amenable to modifications to their physical properties such as density and melting point via the introduction of different chemical functionalities. For these reasons, EILs are being considered as next generation hypergolic propellants.

The most important aspect of a hypergolic fuel is its ignition delay (ID) when mixed with an oxidizer, such as Inhibited Red Fuming Nitric Acid (IRFNA) and White Fuming Nitric Acid (WFNA). The longer the ID is, the larger the combustion chamber must be to avoid pressure spikes that could rupture the engine. A short ID therefore is a necessary condition for hypergolic fuels. The current challenge is to design a hypergolic fuel which is high-performing, nontoxic, safe to handle and transport and has an ID comparable

* Corresponding author; e-mail: dxs@cfdrc.com

to MMH (~5 ms). Therefore, theoretical prediction of IDs in EILs prior to their laboratory synthesis will have a significant impact in reducing time, cost and the risk of failure at later stages of hypergolic propellant development.

Almost all previous research on ignition has been with molecular liquids (tertiary amines and amine azides), although recent work by the Air Force Research Laboratory (AFRL) has initiated studies to understand the ID mechanism of EILs. Recently one of us has attempted[6] to develop a correlation between the ionization potential (IP) and ID for amines. The first step for hypergolic reaction is thought to be electron donation from the amine nitrogen to the NO_2 of IRFNA[7]. Although this correlation (large IP means long ID) worked quite well within the family of tertiary amines and other acyclic alkyl amines (methyl, dimethyl and trimethyl amines) and hydrazine derivatives (MMH and hydrazine), we have been unable to establish correlations amongst members from two different families. Among other related work, Koch[8] attempted to correlate ID to the basicity of alkyl amines using the Hard Soft Acid Base principle (large basicity means short ID). However, results from several other studies reported in the literature contradict these findings. For example, monomethylaminoethylazide (MMAZ, pKa=9.3) has longer ID compared to dimethylaminoethylazide (DMAZ, pKa=8.5). Higher pKa values indicate higher basicity. McQuaid[9] has correlated ID of four tertiary amines with the percent alignment of the amine nitrogen lone pair with respect to the adjacent C-C and/or C-N bond using *ab initio* quantum chemistry calculations. Since the percent alignment is essentially an indirect measure of the basicity of the amine nitrogen, such a correlation may have been fortuitous. Although all earlier works pertained to molecular liquids, it can be inferred that ID is not just a function of a single parameter. Therefore, in order to develop a successful correlation, multiple parameters must be considered.

Attempts to establish correlations between ID times of EILs and their appropriate molecular parameters are reported here for the first time. We follow two approaches: 1) Quantitative Structure Property Relationship (QSPR), and 2) Artificial Neural Networks (ANNs). A large number of descriptors of various types, such as thermodynamic, constitutional, charged partial charged surface area (CPSA) and electrostatic potentials were considered for the correlation. An automatic procedure was followed to select the best descriptors. ANNs are inherently non-linear in nature (as opposed to QSPR, which is linear), so the selected descriptors were further used to examine which of the two approaches was more suitable for ID correlations.

RESULTS AND DISCUSSION

QSPR PREDICTIONS OF IGNITION DELAY

Since QSPR is a statistical method, collecting a large number of ID data is important. In this work, data were obtained from the literature and private communications. A total of 27 IDs for ionic liquids (ILs) with WFNA were collected.

Figure 1 shows the cations and anions of the ILs and Table 1 shows their composition and IDs. The QSPR relates a particular property to a set of descriptors using the following linear equation:

$$P = C_0 + \sum C_i D_i$$

where P is the property, D_i*'s* are the descriptors and C_i*'s* are the adjustable coefficients of the descriptors. One of the most important aspects of QSPR is to find meaningful descriptors that describe the property with high correlation. In this work, a large number of descriptors were calculated, primarily using the quantum semi-empirical method, PM3[10]. This technique is one of the most widely used approaches for quick calculations of molecular properties of large systems. PM3 predictions are known to provide more than 38% improvements over the AM1 semi-empirical method for 657 molecules containing variety of atoms and functionalities. The molecular geometries of all ILs were optimized using PM3 semi-empirical Hamiltonian before computing the descriptors.

Descriptor Selection

ID has a complex dependence on propellant chemical reactions and mixing. From quantum mechanical point of view, molecular reactivity originates from the nature of charge/electron distribution around the molecule. Moreover, mixing depends on bulk properties, such as surface tension. It has been shown by Stanton and Jurs[11] that surface tension and boiling points can be well correlated with descriptors related to the charged partial charge surface area (CPSA) of molecules. Therefore the descriptors derived from PSA are expected to be suitable for ID correlation as well. Politzer's group[12] has shown that some electrostatic potential (V_S^+ , V_S^-, σ_{tot}^2, ν) derived descriptors can describe properties such as the impact sensitivity[13], heat of sublimation and vaporization[14], heat of crystallization[15], diffusion coefficient[16] and solubility[17, 18]. Therefore, we have included these electrostatic descriptors in our descriptor list. We have also selected many thermodynamic descriptors, which include heat of combustion, heat of formation, and specific heat. Lastly, we have included some constitutional descriptors, such as actual and relative numbers of H, N and O atoms in the ILs. Over 160 descriptors, most of them of the PSA type, have been examined to establish a correlation. These descriptors were computed using the Codessa Pro code[19].

Descriptor Fitting

Once the descriptors were selected, multi-linear regression was performed. Two ID data points (for S16 and S23) were eliminated from the data set as their values appeared to be disconnected from the other ILs. In order to have a statistically better correlation, more data points between S22 and S16, and S16 and S23 are required with sufficient spread. For this large data set of descriptors, an automated selection procedure was used, as implemented in CodessaPro[19] to choose two to five descriptors that correlated well with the ID data. The rule of thumb is that the number of descriptors chosen should be within 1/6 to 1/3 of the number of data points. The selection was based on the correlation coefficient, R^2. A large correlation coefficient indicates a better correlation fit. In addition to correlation coefficients, automatic cross validations were also performed in order to obtain values for the cross-validation coefficient, R^2cv. This was performed by the commonly used "leave-one-out" method, according to which one of the data points is left out before doing the correlation fit, and later it is used for validation purposes. This procedure was repeated for all the data points. Large R^2cv indicates better predictive ability of the fit. Although R^2cv is generally lower than R^2, a R^2cv within 25% of R^2 is considered good.

Results of QSPR Fitting

The automatic procedure for establishing a QSPR resulted in five best descriptors for the 25 ID data points, as shown below:

1. $\sqrt{\nu\sigma_{tot}^2}$ (electrostatic potential derived descriptor as reported by Politzer)
2. n_H = number of H atoms,
3. FPSA = (atomic charge weighted positively charged surface area)/(total molecular surface area),
4. RNCS = relative negative charged surface area = SA_{MNEG} x RNCG, where SA_{MNEG} is the surface area of the maximum negatively charged atom, and RNCG = (charge of most negative atom)/(sum total negative charge),
5. SA_N = surface area for nitrogen atoms

All charged and electrostatic potential derived descriptors were calculated using the semi-empirical PM3 Hamiltonian while the surface areas were evaluated using the Connolly algorithm. It is intriguing that the CPSA descriptors found here have also been the most appropriate descriptors found by Stanton and Jurs[11] for the prediction of surface tension and boiling points. It is also interesting to see that none of the thermodynamic and constitutional descriptors (except n_H, the number of H atoms) were involved in the best QSPR fitting.

Figure 2 shows the QSPR results with all 25 ILs using the above five descriptors. The plot shows a comparison of experimental and predicted ID values. The data points are nearly uniformly distributed over both sides of the ideal line. The R^2 and R^2cv values are 0.70 and 0.46, respectively. There are four (S2,

S13, S22, and S27) outliers (marked by ovals). Although the correlation is not perfect, it clearly indicates that a correlation does exist between certain descriptors and ID times. Moreover, we expect that with the availability of more data points, it will be possible to obtain a more robust correlation. At this time, experimental uncertainties in the ID measurements are not available, so the goodness of the present correlation fit for these errors could not be assessed.

In the next step, we removed the outliers and observed a significant improvement in the correlation. Figure 3 shows a new comparison of experimental and predicted ID values. The R^2 and R^2cv are 0.91 and 0.83, respectively. We also expect that this correlation can be improved further once IDs for more ILs are available. Finally, the following two correlations were obtained:

Correlation with 25 ILs:

$$ID = -810.374 + 0.162209(\sqrt{\nu\sigma_{tot}^2}) + 0.53287 \cdot SA_N + 6521.88 \cdot FPSA + 30.8281 \cdot RNCS + 19.3 \cdot n_H$$

Correlation with 21 ILs:

$$ID = -741.626 - 0.236786(\sqrt{\nu\sigma_{tot}^2}) + 0.296784 \cdot SA_N + 5859.16 \cdot FPSA + 36.9699 \cdot RNCS + 20.5 \cdot n_H$$

The QSPR assumes a linear relationship between the descriptors and ID times. However, since the process of ignition for an IL is very complex, it is likely that this relationship may be nonlinear. In order to examine this, we trained various Artificial Neural Networks (which are inherently nonlinear in nature) and used them for prediction analyses. In this method we did not remove the outliers.

ANN PREDICTION OF IGNITION DELAY

Artificial Neural Networks (ANNs) are mathematical constructs that can be trained to learn relationships between a set of factors and the results, even if the relationship between them is theoretically undefined and is non-linear[20]. The concept originated from attempts to model the functioning of the human brain. Subsequently, neural networks have been successfully applied in various fields of chemistry, such as structure-activity relationships, structure-spectrum correlations, catalyst activity predictions [21-23], chemistry kinetics, [24, 25] and controls[26]. Like QSPR, the ANN approach is data driven and ANNs can be trained to predict an output (ignition delay) from a given set of inputs (descriptors). One of the most common networks, which we have used here, is the feed-forward Multi Layered Perception (MLP) ANN. The MLP consists of input layer containing units representing possible descriptors controlling ignition delay, an arbitrary number of hidden layers, and an output layer containing units representing ignition delay, as shown in Figures 4a and b. Commonly, instead of the bias neurons (blue), signal thresholds are added as summing junctions to ensure that the ANN can generate a non-zero neural output for a zero input, regardless of the value of the weights. The neuron models are joined together by adaptive connections in a feed-forward data flow configuration. The values of the connections are typically adjusted using ANN "training" algorithms, so that a fully trained ANN can produce desired outputs for given inputs. ANNs are regarded as universal approximations: an MLP with a single hidden layer can approximate any continuous function to any degree of accuracy [27].

Training Procedure

The training procedure is based on iterative adjustments of the internal weights in response to the error between the desired and the predicted ANN output. The approach is illustrated in Figure 4c. The steps involved in ANN configuration and training are as follows:

1. Generate ANN input/output data consisting of sets of appropriate descriptors for the experimental IDs of the ILs.
2. Scale the data (if necessary) using appropriate scaling techniques, and normalize it, typically in ranges (0, +1) or (-1, +1).

3. Define the network topology in terms of the number of layers and the number of neurons per layer. Select an appropriate transfer function and initialize the network with small, random weights. Several trials are required to obtain a favorable topology for ID prediction.
4. Feed the input (i.e. descriptors) and target data (experimental IDs) into the network once Steps 1-3 are completed. ANN training is accomplished using the supervised learning approach [20].
5. Calculate the ANN output (predicted ID). Calculate the training error based on the RMS difference between the calculated and the target IDs (experimental) and adjust the weights using an error optimizer. In this work, the gradient descent Scaled Conjugate Gradient [28] (SCG) optimizer was used (as outlined below) to iteratively adjust the weights based on training error minimization.
6. Repeat Steps 4 and 5 until the error was below a threshold level.

Scaled Conjugate Gradient

The training algorithm involves error correction learning, and is based on the comparison of the ANN output to its target signal. The objective is to minimize a cost function, E, defined as a measure of the error between the predicted and target signals. Of the several measures available, the most common is the L^2 norm defined by:

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{j}^{outputs} e_j^2(\boldsymbol{w}) \equiv \|E\|^2; \qquad e_j(\boldsymbol{w}) = d_j - y_j(\boldsymbol{w})$$

where d_j is the target value of every output neuron y_j. Network training thus involves repetitive adjustment of the weight $\boldsymbol{w}$ until the error requirements are satisfied, or until termination.

Minimization of the cost function E($\boldsymbol{w}$) is a problem of optimization. Most of the optimization techniques used for training of neural nets, including the SCG used here, are based on gradient descent approximations. Accordingly, given a weight vector w^t at iteration *t,* the estimate of its value at *t + 1* is given by a Taylor expansion of the form:

$$w^{t+1} = w^t + \eta \boldsymbol{g}; \qquad \boldsymbol{g} = -\Delta_w E^t; \eta > 0$$

where $\Delta_{\boldsymbol{w}} E^t$ is a measure of the gradient of E^t with respect to w^t. The SCG approach was used in a marching algorithm to find $\boldsymbol{w}$ which minimizes *E*; in this case η specifies the step size and the gradient $\boldsymbol{g}$ determines the marching direction in weight space. A general form of such an algorithm is given by pseudo-code [28]. As shown below, weight update consists of two steps: calculation of the vector $\boldsymbol{p}$ pointing in the direction of minimum *E* in weight space, and calculation of step size η in that direction. Details are given by Moller [28].

1. Set *t* = 1 and initialize weight vector $\boldsymbol{w}^t$
2. Calculate the direction $\boldsymbol{p}^t$ and step size $\boldsymbol{\eta}^t$ such that $E\left(\boldsymbol{w}^t + \eta^t \boldsymbol{p}^t\right) < E\left(\boldsymbol{w}^t\right)$
3. Update $\boldsymbol{w}^{t+1} = \boldsymbol{w}^t + \eta^t \boldsymbol{p}^t$
4. If $\Delta_{\boldsymbol{w}} E = 0$, return $\boldsymbol{w}^{t+1}$; else *t* = *t+1*, and go to Step 2.

Training/Validation Procedure

Although a network can be trained to within a very small training error, the latter however does not guarantee that the prediction error from this network will also be of the same order. Generally, the prediction error tends to increase when the training error becomes too low. In that case, the network becomes over trained and tends to memorize the data, thus reducing its generality (or predictive) capabilities [20]. To optimize ANNs' capabilities we used the k-fold validation procedure, as shown in the following five steps:

1. The total of 25 ID data points were divided randomly into 5 groups.
2. For each group, 20 ID data points were used for training and the remaining 5 used for validation. The data were arranged so that no two validation sets were the same. This training/validation procedure was repeated for each group (for a total of 5 training sessions) resulting in a 5-fold validation.
3. Effective weights (discussed below) were determined from the training sessions.
4. Final weights were obtained for the case in which the prediction error was minimized.

ANN Results

The quality of ANN training typically depends on (a) ANN topology, and (b) data preparation. The predictive capabilities depend on evaluation of the final effective weight vector obtained from the weights determined in the individual training sessions in the k-fold validation process.

a) ANN Topology

Topology generally refers to the ANN layout. An *i-m-n-o* topology designation means that the ANN contains i number of inputs (i.e. 5 descriptors for each IL), o number of outputs (i.e. one ID for each IL) and 2 hidden layers with m and n number of neurons in each layer. It should be noted that there is no rule of thumb for topology selection and the most appropriate topology is typically selected via trial and error. To determine the optimal layout, we performed numerical experiments using the following topologies: 5-5-1 (36 weights), 5-10-1 (71 weights), 5-10-5-1 (121 weights), and 5-10-10-1 (181 weights). We found that the topology with 2 hidden layers containing 10 and 5 neurons (121 weights) was the most appropriate, as the 5-fold cross validation produced the least error with this configuration. Most of our calculations were performed using a "sigmoid" transfer function. The bias was set to -0.5.

b) Data Preparation

ANN training characteristics can exhibit sensitivity to data normalization and scaling.

Normalization: Input and output data normalization is an important aspect of training the network and performed to avoid problems with saturation of the neuron transfer function. Input and output data are typically normalized in the range (0,1) or (-1,+1). The type of normalization is problem dependent and may have some effects on how well the ANN trains. Typical procedure for normalization, for example, in the (0,1) range is: $y^{norm} = (y - y_{min})/(y_{max} – y_{min})$, where y is the set of data points. The minimum and maximum values are found from the data set. We initially performed the data normalization in two ways:

i. *Full Data*: The min-max values were taken from the whole data set that contained both the training and validation data points. In this case all of the validation data points were within the min-max values and properly normalized.

ii. *Local (Partial) Data*: The min-max values were taken from the training data set that was extracted from the whole data set and therefore was a sub-set. Such data sets were generated from k-fold cross-validation procedures. In order to insure that the validation data set was within the min-max range, we extend the range of the latter using a user-defined safety parameter, s^{range}, as follows: $y_{min(safe)} = y_{min}(1- s^{range})$; $y_{max(safe)} = y_{max}(1+ s^{range})$. The same procedure can also be used with the *Full Data* method. In all of the simulations performed here, we used $s^{range}=0.5$ for normalizing local data.

Scaling: If the data set spans a large range of values, it may be beneficial to scale it prior to normalization. For positive-valued data, such scaling is typically done with a logarithmic function of the form: $y^{scaled} = \log(y)$. The scaled data can then be normalized in the range (0,1) or (-1,+1). Table 2 shows the different types of normalization and scaling tried in this work. Figure 5 and 6 show the prediction plots for one of the 5-fold cross validation cases. Clearly for one data set, as the training error increases, the prediction error (errRMS) increases. Prediction errors appeared to be better with training errors approximately between 0.01 and 0.1. In addition the , "*full data minMax (-1,+1) range"* and "*local data minMax (0, +1) range*" normalization procedures produced even better fitting than (i) and (ii). We

therefore continued rest of the cross validation calculations with 5-10-5-1 (121 weights) and these two data normalization procedures.

We adopted two techniques for computing the weight vector valid for the whole data set (25 points).

Weight Averaging: In this approach we trained the ANNs separately for each of the five data sets in the 5-fold validation process, resulting in five different weight vectors for the selected ANNs topologies and data preparation approaches. The final weights were calculated as linear averages of weights from the 5 cross validation calculations. These weights were then used in the network to compute IDs of all 25 ILs. Figure 7 shows the ANN predictions using the two different data normalizations with a training error of 0.1. Although we obtained a trend, the predictions are clearly not satisfactory. In the above procedure, for each of the 5-fold cross validation trainings, we used the same initial guess for the weights. We observed that for each of the 5-fold trainings, the converged weights were significantly different. This implies that for each training, the weights converged to different local minima. In order to improve the predictions and obtain a more consistent set of weights, we adapted the Sequential Training approach described below.

Sequential Training: In this approach, we used the weights of the 1st training and used the first data set as the initial values for the 2nd training, and so on for k-1 data set for kth training. We denote this symbolically as 1→ 2→ 3→ 4→ 5 training. For each training session, the ANNs were trained to a chosen error. The final weights at the end of the 5th cross validation training were then used for predictions of all 25 ILs. In order to check for training bias, we repeated the 1→ 2→ 3→ 4→ 5 training (called "Restart 1-5") using 5→ 4→ 3→ 2→ 1 training ("Restart 5-1"). Figure 8 shows the final results of ANN predictions of IDs of all 25 ILs. An excellent correlation was obtained using the procedure mentioned above, and there was no significant difference in the results between "Restart 1-5" and "Restart 5-1". The weights used for these predictions were taken from the final training in each particular class: final training on data set5 for Restart1-5; and final training on data set1 for Restart5-1. It should be pointed out that in the QSPR work, reported earlier, four ILs out of twenty five were found to be outliers, and they were eliminated to obtain a better correlation. However, in the ANN work reported here, all 25 ILs were included. We thus suspect that ID is not entirely a linear function of the descriptors chosen. Since ANN is inherently non-linear in nature, it provides a better correlation between the IDs and the descriptors.

SUMMARY AND CONCLUSIONS

This work attempts to develop Quantitative Structure Property Relationship (QSPR) and Artificial Neural Network (ANN) correlations between predictive and experimental ignition delays of a number of energetic ionic liquids with nitric acid and a set of molecular descriptors. Using the linear QSPR approach, ignition delays were found to correlate reasonably well with the five descriptors once the four outliers were removed from the data set. The five descriptors were chosen from a pool of more than 160, and three of the five descriptors were based on partial charged surface area. Since ignition of ionic liquids with WFNA is a very complex phenomenon involving mixing, heat transfer and chemical reactions, we expect that the relationship between these descriptors and the ignition delay is non-linear. As ANNs are inherently non-linear in nature, we examined their performance using the same descriptors as that used in QSPR. Selected ANN topologies, training procedures, data normalizations and scaling techniques were examined. Data predictions were extensively studied with the 5-fold cross validation technique. With successive trainings, we obtained an excellent correlation for all 25 ionic liquids.

ACKNOWLEDGMENT

We gratefully acknowledge the Air Force Research Laboratory for funding this work under SBIR contract # FA8650-09-M-2037.

REFERENCES

[1] S. Schneider; T. Hawkins; M. Rosander; G. Vaghjiani; S. Chambreau; G. Drake, Ionic Liquids as Hypergolic Fuels, *Energy Fuels* **2008,** *22*, (4), 2871

[2] H. Gao; Y.-H. Joo; B. Twamley; Z. Zhou; J. n. M. Shreeve, Hypergolic Ionic Liquids with the 2,2-Dialkyltriazanium Cation, *Angew. Chem. Int. Ed.* **2009,** *48*, 2792.

[3] Y. Zhang; H. Gao; Y. Guo; Y.-H. Joo; J. n. M. Shreeve, Hypergolic N,N-Dimethylhydrazinium Ionic Liquids, *Chem. Eur. J.* **2009** *(in press)*.

[4] L. He; G.-H. Tao; D. A. Parrish; J. M. Shreeve, Nitrocyanamide-Based Ionic Liquids and Their Potential Applications as Hypergolic Fuels, *Chem. Eur. J.* **2010** *(in press)*.

[5] S. D. Chambreau; S. Schneider; M. Rosander; T. Hawkins; C. J. Gallegos; M. F. Pastewait; G. L. Vaghjiani, Fourier Transform Infrared Studies in Hypergolic Ignition of Ionic Liquids, *J. Phys. Chem. A* **2008,** *112*, (34), 7816.

[6] D. Sengupta; S. Raman, Development of Improved Performance, Low Toxic Hypergolic Fuel for Replacing Monomethylhydrazine, *Phase I Technical Report, Army Aviation and Missile Command, Redstone Arsenal Contract # W31PQ06C0167*.

[7] I. Frank; A. Hammerl; T. M. Klapoetke; C. Nonnenberg; H. Zewen, Process During the Hypergolic Ignition Between Monomethyhydrazine(MMH) and Nitrogen Tetroxide (N2O4) in Rocket Engines, *Propellant, Explosives, Pyrotechnics* **2005,** ***30***, 44.

[8] E.-C. Koch, Acid-base interaction in energetic materials:I. The hard and soft acids and bases (HSAB) principle – Insight to reactivity and sensitivity of energetic materials, *Propellants, Explosives, Pyrotechnics* **2005,** *30*, 5.

[9] M. J. McQuaid, Computationally based measures of amine azide basicity and their correlation with hypergolic ignition delay, *Army Research Laboratory Report, ARL-TR-3122* **December 2003**.

[10] J. J. P. Stewart, Optimization of parameters for semiempirical methods II. Applications, *J. Comput. Chem.* **1989,** *10*, 221.

[11] D. T. Stanton; P. C. Jurs, Development and Use of Charged Partial Surface Area Structural Descriptors in Computer-Assisted Quantitative Structure-Property Relationship Studies, *Anal. Chem.* **1990,** *62*, 2323.

[12] P. Politzer; J. S. Murray, The fundamental nature and role of the electrostatic potential in atoms and molecules, *Theor. Chem. Acc.* **2002,** *108*, 134.

[13] J. S. Murray; P. Lane; P. Politzer, Effects of strongly electron-attracting components on molecular surface electrostatic potentials: application to predicting impact sensitivities of energetic molecules, *Molecular Physics* **1998,** *93*, 187.

[14] P. Politzer; J. S. Murray; M. E. Grice; M. Desalvo; E. Miller, Calculation of heats of sublimation and solid phase heats of formation, *Molecular Physics* **1997,** *91*, 923.

[15] P. Politzer; J. S. Murray, Relationships between Lattice Energies and Surface Electrostatic Potentials and Areas of Anions, *J. Phys. Chem. A* **1998,** *102*, 1018.

[16] P. Politzer; J. S. Murray; P. Flodmark, Relationship between Measured Diffusion Coefficients and Calculated Molecular Surface Properties, *J. Phys. Chem.* **1996,** *100*, 5538.

[17] J. S. Murray; S. G. Gagarin; P. Politzer, Representation of C_{60} Solubilities in Terms of Computed Molecular Surface Electrostatic Potentials and Areas, *J. Phys. Chem.* **1995,** *99*, 12081.

[18] P. Politzer; J. S. Murray; P. Lane; T. Brinck, Relationship between solute molecular properties and solubility in supercritical CO_2, *J. Phys. Chem.* **1993,** *97*, 729.

[19] A. R. Katritzky; K. Mati; P. Ruslan, Codessa Pro, Comprehensive Descriptors for Structural and Statistical Analysis, CompuDrug International, Inc., 115 Morgan Drive, Sedona, AZ 86351, USA.

[20] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd Edition, Upper Saddle River, NJ, Prentice Hall Inc (1999).

[21] T. Hattori; S. Kito, Neural network as a tool for catalyst development, *Catalysis Today* **1995,** *23*, 347.

[22] U. Rodemerck; M. Baerns; M. Holena; D. Wolf, Application of a genetic algorithm and a neural network for the discovery and optimization of new solid catalytic materials *Applied Surface Science* **2004,** *223* 168.

[23] J. Zupan; J. Gasteiger, Neural networks: A new method for solving chemical problems or just a passing phase? *Analytica Chimica Acta* **1991,** *248*, 1.

[24] J. A. Blasco; N. Fueyo; C. Dopazo; J. Ballester, Modelling the Temporal Evolution of a Reduced Combustion Chemical System with an Artificial Neural Network *Combust and Flame* **1998,** *113*, 38.
[25] M. Z. Pindera, Intelligent Solution Optimization Using Artificial Neural Networks Applied to Numerical Flow Solvers, *NASA SBIR Phase I Final Report, NASA Contract Number: NAS2-98053* **(Oct. 1998)**.
[26] M. Z. Pindera, Adaptive Flow Control Strategies Using Simple Artificial Neural Networks, *AIAA-2002-0990* **2002**.
[27] K. Hornik; M. Stinchcombe; H. White, Multilayer Feedforward Networks are Universal Approximations, *Neural Networks* **1989,** *2*, xxx
[28] M. F. Møller, A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning." *Neural Networks* **1991,** *6*, 525.

TABLES AND FIGURES

Table 1 : Ionic liquids considered here and their ignition delay (ID) times with WFNA in milliseconds.

Ionic Liquid	Cation	Anion	ID (ms)
S1	2	NCA	16
S2	2	DCA	22
S3	2	NO_3-	4
S4	1 (R= allyl	DCA	43
S5	1(R=propargyl)	DCA	14
S6	1 (R=NH_2)	DCA	31
S7	1-R (R=butyl)	DCA	47
S8	1-butyl-1-methyl-pyrrolidinium	DCA	44
S9	N-butyl-3-methylpyridinium	DCA	37
S10	3 (R1=Me)	DCA	58
S11	3 (R1=C_2H_5)	DCA	22
S12	3 (R1=C_4H_9)	DCA	46
S13	3 (R1=CH_2CHCH_2)	DCA	24
S14	3 (R1=CH_2CCH)	DCA	30
S15	3 (R1=CH_2CH_2OH)	DCA	40
S16	3 (R1=CH_2CN)	DCA	1286
S17	3 (R1=Me)	NCA	125
S18	3 (R1=C_2H_5)	NCA	198
S19	3 (R1=C_4H_9)	NCA	228
S20	3 (R1=CH_2CHCH_2)	NCA	130
S21	3 (R1=CH_2CCH)	NCA	134
S22	3 (R1=CH_2CH_2OH)	NCA	247
S23	3 (R1=CH_2CN)	NCA	1642
S24	1(R=ethyl)	NCA	46
S25	1 (R=n-butyl)	NCA	78
S26	1 (R=allyl)	NCA	81
S27	1 (R=cyanomethyl)	NCA	65

Table 2: Various data normalizations investigated in this work.

local data minMax (0,+1) range	minMax values required for data normalization based on local sampled data of 20 points. To encompass the values of the validation data set, min. value was reduced by 50% and max. value was increased by 50%. Data were then normalized in the interval (0, +1).
full data minMax (0,1) range	Values required for data normalization based on complete data set of 25 points. This will include values of any validation data automatically. Data were then normalized in the interval (0, +1).
full data minMax (-1,+1) range	Values required for data normalization based on complete data set of 25 points. This will include values of any validation data automatically. Data were then normalized in the interval (-1, +1).
full scaled data minMax (0,+1) range	Full data were first scaled by a logarithmic function and then minMax values were found. Data were then normalized in the interval (0, +1). Such scaling is useful when the I/O covers a large range of values.
full scaled data minMax =>> (hTan squashing)	This is a reference to the internal structure of a neuron. Here it transforms the data using a hyperbolic tangent function. This helps in maintaining the neuron response in the linear range.

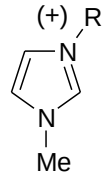
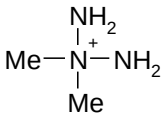
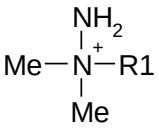
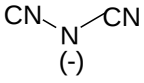
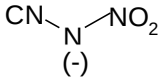
 **1**	 **2**	 **3**
 **DCA**	 **NCA**	

Figure 1: Cations and anions that constitute the ionic liquids considered here.

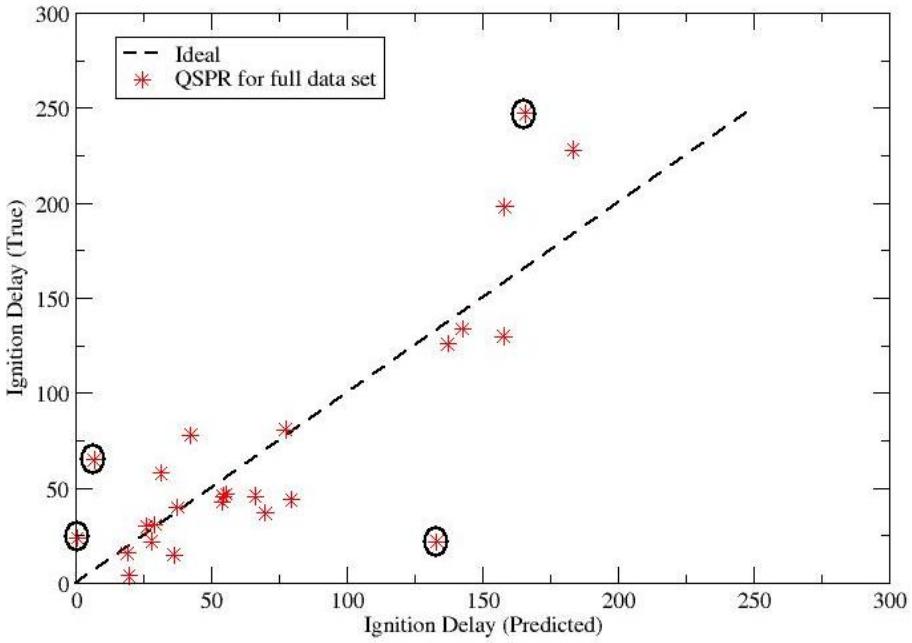


Figure 2: QSPR for ID predictions. 25 ILs have been included in the data set. Outliers are marked by ovals.

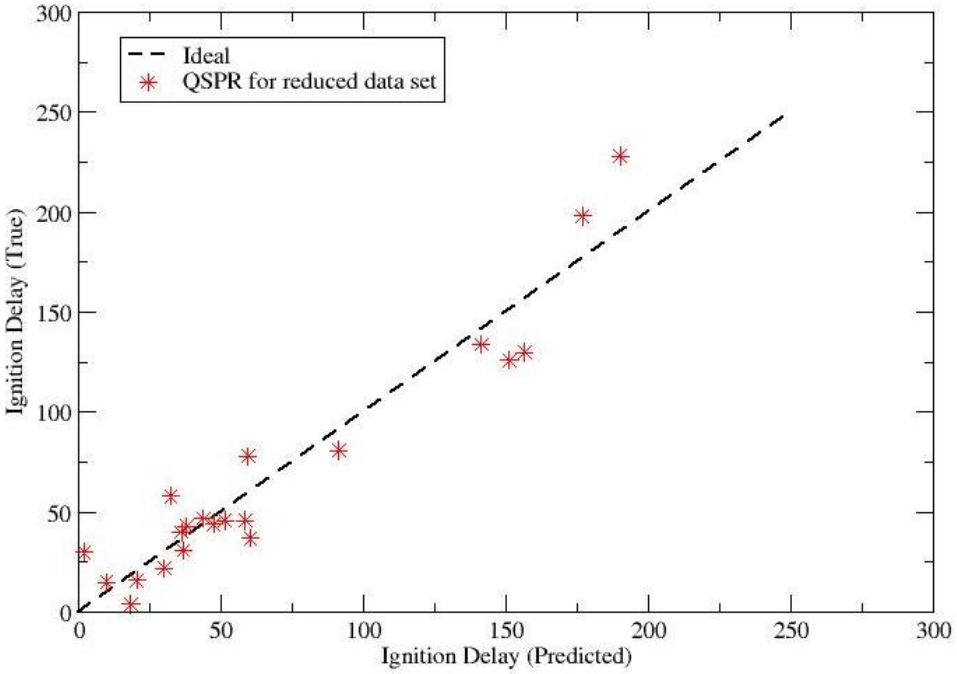
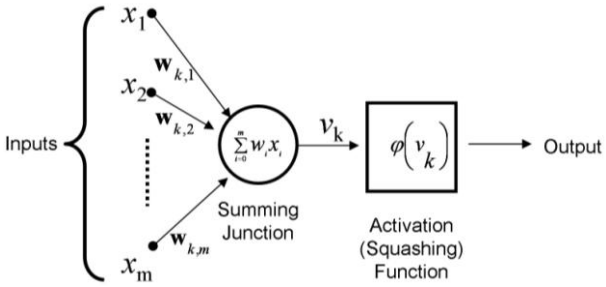
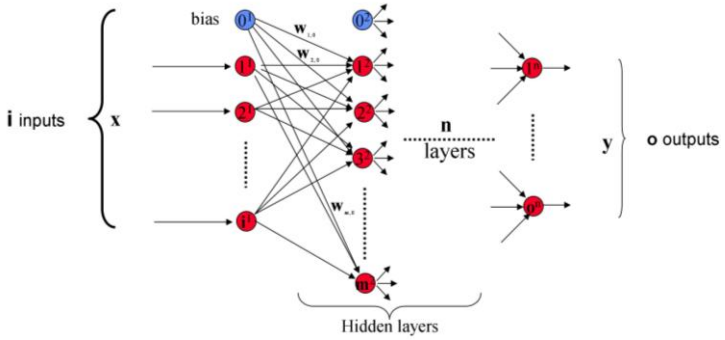


Figure 3: QSPR for ID predictions. Four outliers from Figure 2 have been excluded.

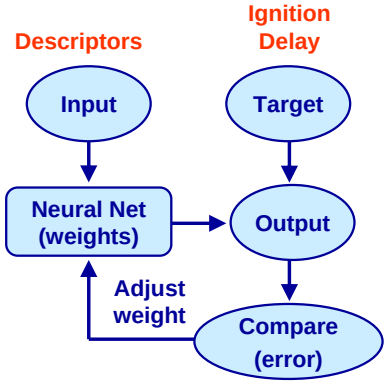


$$y_k = \varphi\left(\sum_{i=0}^{M} w_i x_i\right)$$

(a)



(b)



(c)

Figure 4: a) an example of a neuron model; b) an example of neural network: φ, w, x and y are activation functions, weights, inputs and outputs, respectively. Common activation functions include the sigmoid ($\varphi = 1/\left[(1+\exp(-\alpha v))\right]$); with output in the (0,+1) range, or the hyperbolic tan ($\varphi = \left[\exp(\alpha v) - \exp(-\alpha v)\right]/\left[\exp(\alpha v) + \exp(-\alpha v)\right]$); with output in the (-1,+1) range; c) Neural network training procedure.

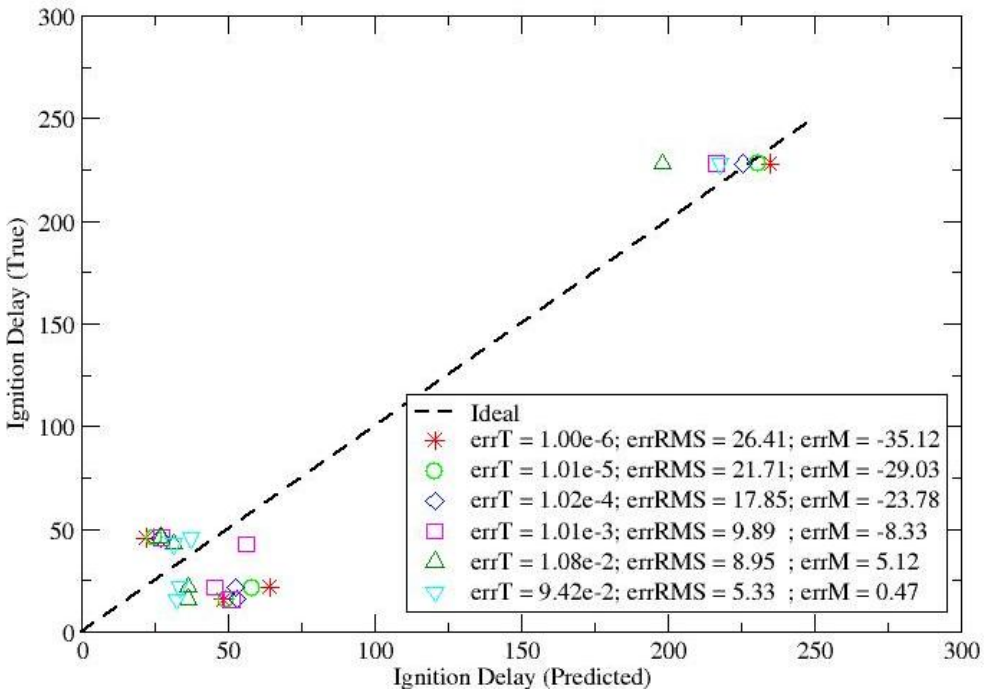


Figure 5: Predictions at different training errors (errT) with 121 weights and "*full data minMax (-1,+1) range"* normalization. As the training error decreases the prediction error increases. This plot is for one of the 5-fold cross validation cases. 20 data points were used for training and 5 for validation.

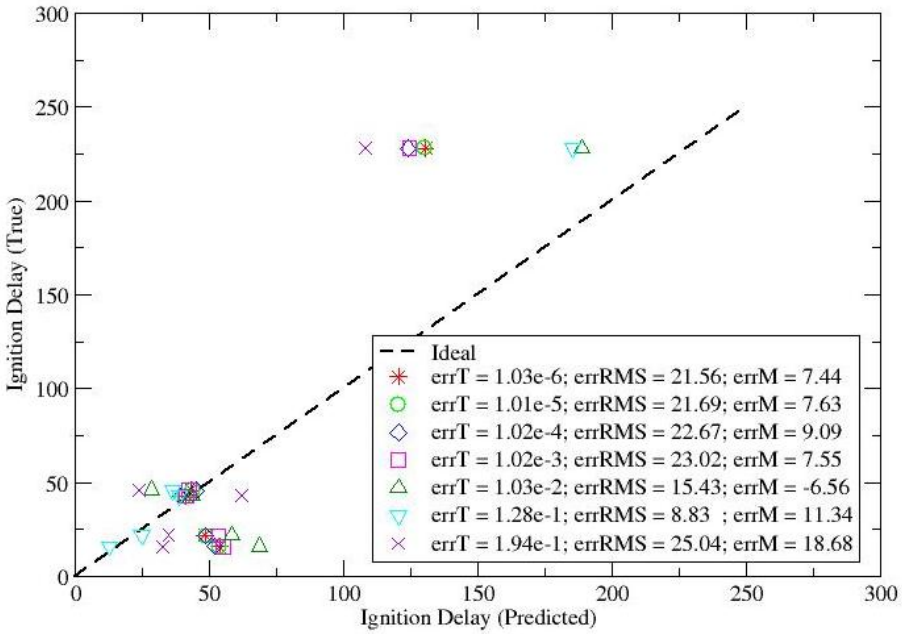
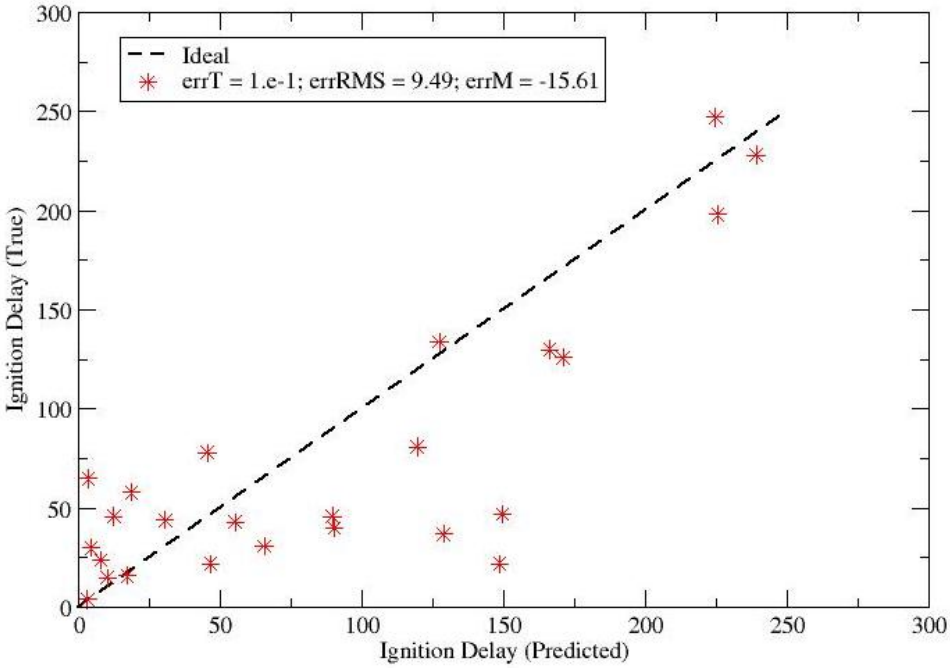
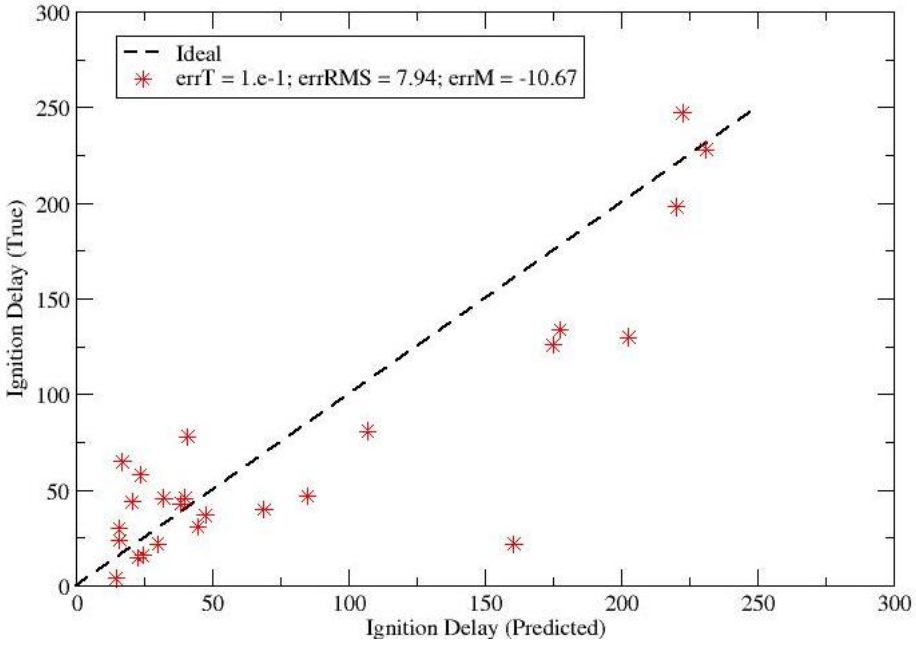


Figure 6: Predictions at different training errors (errT) with 121 weights and *"local data minMax (0, +1) range"* normalization. As the training error decreases the prediction error increases. This plot is for one of the 5-fold cross validation cases. 20 data points were used for training and 5 for validation.

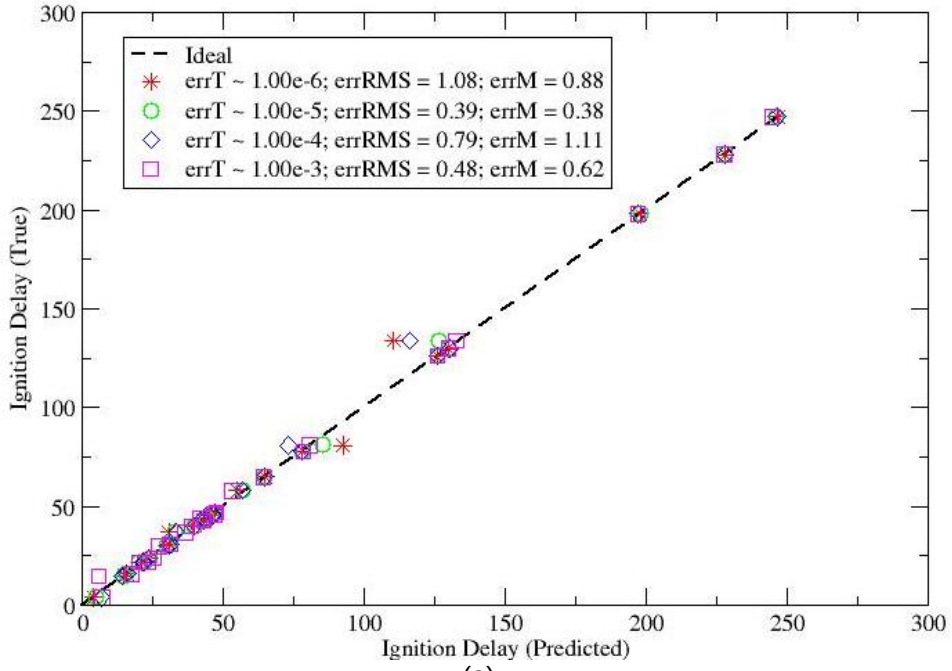


(a)

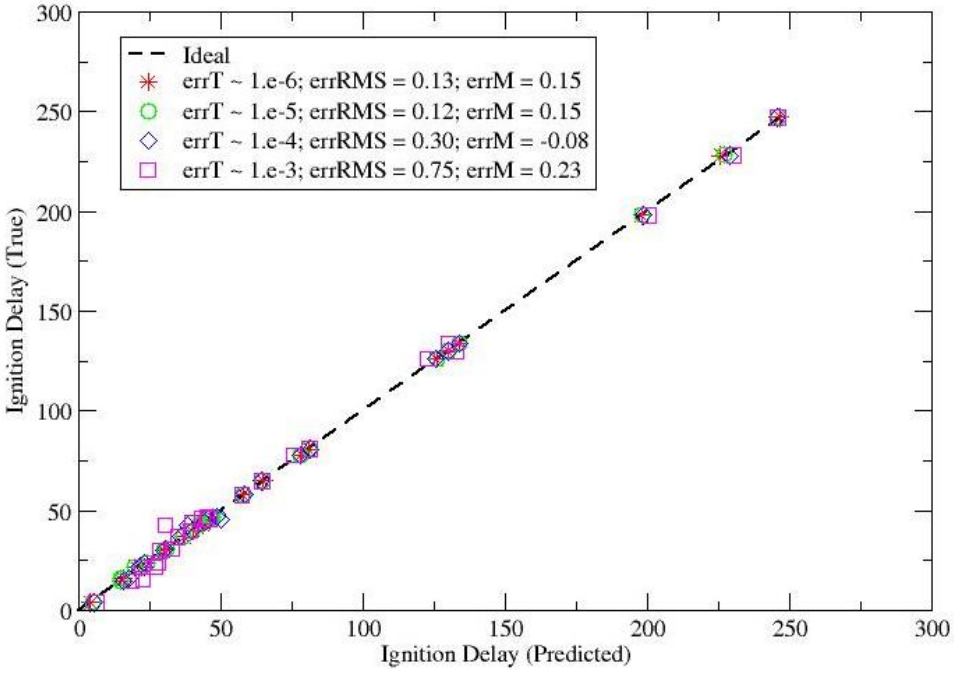


(b)

Figure 7: ID predictions with the two normalization approaches: (a) *"local data minMax (0, +1) range"*, (b) *"full data minMax (-1, +1) range"*.



(a)



(b)

Figure 8: Final ANN predictions of IDs of all 25 ILs using *"full data minMax (-1,+1) range"*: (a) Restart 1-5 and (b) Restart 5-1. It should be noted that predictions are excellent at all training errors.